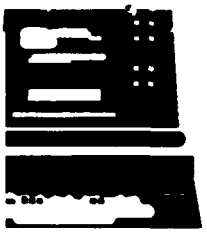


USAISEC

US Army Information Systems Engineering Command
Fort Huachuca, AZ 85613-5300

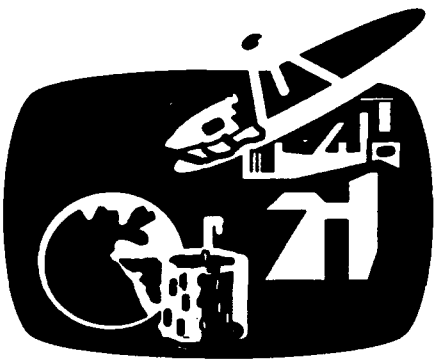
U.S. ARMY INSTITUTE FOR RESEARCH
IN MANAGEMENT INFORMATION,
COMMUNICATIONS, AND COMPUTER SCIENCES
(AIRMICS)

Asymptotic Performance Analysis of Hybrid ARQ Protocols in Slotted Direct-Sequence Code-Division Multiple-Access Networks

(ASQB-GC-90-020)

16 March 1990

115 O'Keefe Bldg
Georgia Institute of Technology
Atlanta, GA 30332-0800



This paper examines the steady-state and dynamic performance of Type 1 Hybrid ARQ protocols in slotted direct-sequence code-division multiple-access networks. The network consists of an arbitrary number of transceivers arranged in a paired-off topology. A Markov model is used to derive the throughput-delay expressions in terms of the channel cutoff rate and capacity. Network design parameters are identified and their dependency on system parameters is examined in detail. It is shown that, for a given population size. traffic intensity, and bit energy-to-background noise ratio, there is an optimal probability of retransmission, code rate, and processing gain that maximizes network performance.

THIS REPORT HAS BEEN REVIEWED AND IS APPROVED

S/ _____

John W. Gowens
Division Chief
CNSD

S/ _____

John R. Mitchell
Director
AIRMICS

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188
Exp. Date: Jun 30, 1986

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS NONE
2a. SECURITY CLASSIFICAION AUTHORITY N/A	3. DISTRIBUTION/AVAILIBILTY OF REPORT N/A
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE N/A	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S) N/A
6a. NAME OF PERFORMING ORGANIZATION AIRMICS 6b. OFFICE SYMBOL *(If applicable)* ASQB-GC	7a. NAME OF MONITORING ORGANIZATION N/A
6c. ADDRESS (City, State, and Zip Code) 115 O'Keefe Bldg. Georgia Institute of Technology Atlanta, GA 30332-0800	7b. ADDRESS *(City, State, and ZIP Code)* N/A
8b. NAME OF FUNDING/SPONSORING ORGANIZATION AIRMICS 8b. OFFICE SYMBOL *(If applicable)* ASQB-GC	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N/A
8c. ADDRESS *(City, State, and ZIP Code)* 115 O'Keefe Bldg. Georgia Institute of Technology Atlanta, GA 30332-0800	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT ACCESSION NO.

11. TITLE *(Include Security Classification)*
Asymptotic Performance Analysis of Hybrid ARQ Protocols in Slotted Direct-Sequence Code-Division Multiple-Access Networks

12. PERSONAL AUTHOR(S)
LTC Joseph M. Hanratty

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT *(Year, Month, Day)*	15. PAGE COUNT
Research	FROM 01/88 TO 03/90	90,03,16	42

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS *(Continue on reverse if necessary and identify by block number)*
FIELD	GROUP	SUBGROUP	Spread Spectrum Networks Code Division Multiple Access (CDMA) Automatic Repeat Request (ARQ) Hybrid ARQ

19. ABSTRACT *(Continue on reverse if necessary and identify by block number)*

This paper examines the steady-state and dynamic performance of Type 1 Hybrid ARQ protocols in slotted direct-sequence code-division multiple-access networks. The network consists of an arbitrary number of transceivers arranged in a paired-off topology. A Markov model is used to derive the throughput-delay expressions in terms of the channel cutoff rate and capacity. Network design parameters are identified and their dependency on system parameters is examined in detail. It is shown that, for a given population size. traffic intensity, and bit energy-to-background noise ratio, there is an optimal probability of retransmission, code rate, and processing gain that maximizes network performance.

20. DISTRIBUTIONN/AVAILIBILTY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
[x] UNCLASSIFIED/UNLIMITED [ ] SAME AS RPT. [x] DTIC USERS	Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE *(Include Area Code)*	22c. OFFICE SYMBOL
LTC Joseph M. Hanratty	(404) 894-3136	ASQB-GC

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete. — SECURITY CLASSIFICATION OF THIS PAGE

Asymptotic Performance Analysis of Hybrid ARQ Protocols in Slotted Direct-Sequence Code Division Multiple-Access Networks

Joseph M. Hanratty and Gordon L. Stüber
School of Electrical Engineering
Georgia Institute of Technology
Atlanta, Georgia 30332
(404) 894-2923

Abstract

The steady-state and dynamic performance of Hybrid ARQ protocols in slotted direct-sequence code division multiple-access networks is analysed. A Markov model is used to derive throughput-delay expressions in terms of the channel cutoff rate and capacity. Network design parameters are identified and their dependency on systems parameters is examined in detail. It is shown that, for a given number of users, traffic intensity, and bit energy-to-background noise ratio, there is an optimal probability of retransmission, code rate, and processing gain that maximises performance.

I Introduction

Code division multiple-access (CDMA) continues to receive interest as a viable network multiple accessing technique for both civilian and military applications. This interest is prompted by the numerous advantages which are afforded by its use, including decentralized control for uncoordinated access, zero access delay, gradual degradation in performance with increased number of users, increased noise immunity, message privacy, and anti-jam (AJ) and low-probability of intercept (LPI) protection. While most of the work to date has dealt with the physical level issues, a number of papers which treat the issues at the link and network levels have recently appeared in the literature. In [1], Raychaudhuri analyzed the CDMA link performance of a paired-off user topology for several packet generation models. More recently, Polydoros and Silvester [2] developed an analytic framework in order to identify those design parameters which summarize the effect of various network characteristics on link level performance. In a paper related to the present paper, Stüber [3] developed an analytic model for evaluating the performance of slotted ALOHA direct-sequence CDMA networks with forward error correction (FEC). In this paper, an approach similar to [3] is followed to investigate the link level issues associated with the use of Hybrid Automatic-Repeat-Request (ARQ) protocols in slotted direct-sequence CDMA networks.

The key performance measures of random access networks are their throughput and delay characteristics. Hybrid ARQ schemes are often employed to improve these characteristics in terms of increased throughput and network reliability. The purpose of this paper is to develop an analytical framework for evaluating and optimizing the throughput-delay performance of Hybrid ARQ protocols in slotted direct-sequence CDMA networks. Asymptotic performance bounds are derived in terms of the channel cutoff rate and capacity. These bounds are considered asymptotic because they assume the use of error detecting/correcting codes operating arbitrarily close to the channel cutoff rate or capacity.

An overview of this paper is as follows. In section II, the network model is defined in terms of its physical and link level characteristics. A Markovian model is employed in section III to derive asymptotic throughput-delay expressions in terms of the channel cutoff rate and capacity. In section IV, numerical results are presented to demonstrate the

asymptotic performance of the CDMA network. Extensive analysis is made to investigate the dependency of the probability of retransmission, code rate, and processing gain on system parameters such as the traffic intensity, number of transceivers, and the level of background noise. Although the effects of jammer noise are not considered here, they provide the motivation for this paper and are addressed in another article [4]. The results of this paper form the 'baseline performance' for the jamming analysis in [4]. Operational considerations are presented in section V. In the appendix, the state transition probabilities are derived for the Markov Model of section III.

II Network Model

The network consists of U radio units arranged in a paired-off user topology similar to that discussed in [1,2]. With this topology, $U/2 = N_T = N_R$, where N_T is the maximum possible number of active transmitters, and N_R is the maximum possible number of active receivers. To simplify the notation, the population size will be expressed in terms of N, the maximum possible number of active receiver-transmitter pairs during a given slot period, where $N = N_T = N_R$. With this topology, only one packet may be directed to a given receiver during a channel slot and there is complete receiver availability whenever the corresponding transmitter is active. Therefore, there is no topological competition, and the channel is not a capture channel. The only disruptive interaction between users is multiple-access interference.

Radio units access the network by using direct-sequence CDMA with a standard slotted ALOHA protocol. Information is transmitted in the form of packets, one packet per time slot. Packet flow for the network is shown in Fig. 1. Each of the network's transceivers can be in one of two modes: origination or blocked. In the origination mode, the probability of transmitting a packet in the i^{th} time slot is geometrically distributed with parameter p_o, where p_o is the probability that the user will transmit the packet in the next time slot.

$$\text{Prob } \{ \text{ packet transmission in the } i^{th} \text{ time slot } \} = p_o(1-p_o)^{i-1} \qquad (1)$$

When either multiple-access interference or background noise causes a packet to be received

in error, a user enters the blocked mode. In this mode, the probability of retransmission is also geometrically distributed but with parameter p_r, where p_r is the probability that a user will retransmit its packet in the next time slot. While a user is waiting to retransmit, it is considered blocked or backlogged because it cannot transmit (but may receive) a new packet until the retransmitted packet is received without error. In practice, limits are set on the maximum number of retransmissions allowed for a given packet. Here, we define the traffic intensity as the average number of packets transmitted per slot, and denote the new and retransmitted traffic intensities as ν_o and ν_r, respectively. At the input to the CDMA channel, new packet transmissions combine with packet retransmissions to form the composite channel traffic. This resulting composite channel traffic can be characterized by its intensity and steady-state arrival distribution. The composite traffic intensity ν simply equals the sum of the new and retransmitted traffic intensities ($\nu = \nu_o + \nu_r$). The composite arrival distribution $f_M(l)$ is the steady-state probability distribution for the number of attempted transmissions M in a given time slot. The form of this distribution depends primarily upon the relative values of p_o and p_r, and on the population size N. Significant simplification in its form results when $p_o = p_r = p$, in which case the distribution becomes binomial with parameters p and N. Further simplification occurs when $p_o = p_r \rightarrow 0$ and $N \rightarrow \infty$, in which case the binomial distribution approaches the Poisson distribution with arrival rate equal to the composite traffic intensity ν.

A Type I Hybrid ARQ protocol is used for error control. This protocol combines the forward-error-detecting (FED) capability of plain (conventional) ARQ with the use of forward-error-correcting (FEC) codes to achieve improved network performance. An example of a Type I Hybrid ARQ channel (hereafter referred to as Hybrid ARQ) is shown in Fig. 2. The ARQ portion of this protocol is of the stop-and-wait (SAW) variety where the sender waits for either a positive (ACK) or a negative (NAK) acknowledgement of the packet just sent before sending another. It is assumed that acknowledgements are made instantaneously over a separate, noiseless return channel. With this assumption, plain ARQ is equivalent to the standard slotted narrow-band ALOHA channel access protocol. If a NAK is received, then the sender retransmits the current packet in the next time slot with

probability p_r. Other ARQ schemes such as go-back-N and selective-repeat can be used to achieve higher throughput efficiency but at the expense of increased network complexity. The FEC portion of the protocol is used to combat the effects of poor channel conditions and tends to reduce the number of retransmissions.

For this analysis, we assume that the channel conditions (i.e., background noise power spectral density (PSD) excluding the multiple-access interference) remain constant. Otherwise, the Type I Hybrid ARQ protocol becomes less efficient as the channel error rate decreases because the extra error correction parity bits are wasted. For time-varying channel conditions, it is better to use the Type II Hybrid ARQ protocol which uses only FED when the channel conditions are good (plain ARQ) and uses both FED and FEC when channel conditions are poor [5]. If channel conditions degrade beyond the capabilities of Type II Hybrid ARQ, code combining techniques (adaptive Hybrid ARQ) can be employed to successfully recover the packet [6,7].

Each packet that is transmitted over the Hybrid ARQ channel of Fig. 2 is comprised of one or more code words. Each code word is used for simultaneous error detection and error correction. The CDMA channel portion of Fig. 2 corresponds to the coding channel for the network. For hard decision decoding, the coding channel is a binary symmetric channel having crossover probability P_S and with the following cutoff rate and capacity:

$$R_o = 1 - \log_2\left(1 + \sqrt{4P_S(1-P_S)}\right), \tag{2}$$

$$C = 1 - H[P_S], \tag{3}$$

where $H[P_S]$ is the binary entropy function given by

$$H(P_S) = -P_S \log_2 P_S - (1-P_S)\log_2(1-P_S). \tag{4}$$

Expressions for the probability of symbol error P_S can be obtained by manipulating (2) and (3). For the cutoff rate case, the probability of symbol error is

$$P_S = \frac{(1-\sqrt{1-\alpha^2})}{2}, \tag{5}$$

where

$$\alpha = 2^{1-R_o} - 1, \tag{6}$$

and for the capacity case

$$P_S = H^{-1}[1 - C], \tag{7}$$

where $H^{-1}(\cdot)$ is the inverse operator of $H(\cdot)$. The probability of symbol error P_S depends upon the type of modulation used by the physical links in our system. For this analysis, we consider the use of binary modulation, e.g. BPSK or DPSK, and results are developed for the DPSK case. The uncoded bit error probability in the presence of additive white Gaussian noise (AWGN) is $P_b = f(\lambda_t)$ where λ_t is the bit energy-to-total noise ratio. For DPSK, $f(\lambda_t) = \frac{1}{2}\exp\{-\lambda_t\}$, and for BPSK $f(\lambda_t) = Q(\sqrt{2\lambda_t})$ where

$$Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{\xi^2}{2}} d\xi. \tag{8}$$

When coding is used, the above code symbol error probability becomes

$$P_S = f(r\lambda_t), \tag{9}$$

where r is the code rate.

CDMA is a spread spectrum multiplexing technique which is characterized by the use of high rate (many chips per code symbol) pseudonoise (PN) spreading sequences. Spreading sequences are selected from code families having good (thumbtack) autocorrelation function properties for accurate acquisition, low cross-correlation function properties for reduced multiple-access interference, and membership to a large family size to support a large number of users. It is assumed that the period of the spreading sequence is much greater than the code symbol duration. It is also assumed that the spreading sequences are completely random so that the resulting multiple-access interference can be modeled as a Gaussian random variable. In practice, spreading sequences are pseudorandom, but for certain code families (Gold codes) the Gaussian assumption is valid to within a scale factor [8,9]. For this analysis, a receiver-based spreading sequence distribution is used. To transmit a packet, a sender modulates each packet bit with the spreading sequence of the intended receiver.

The following simplifying assumptions are made:

1. The usual bursty user assumption ($p_r > p_o$) is relaxed to allow these probabilities to vary freely over their entire range of values ($0 \leq p_r, p_o \leq 1$).

2. The traffic characteristics are homogeneous throughout the network. Therefore, the probability distribution of the multiple-access interference is the same at each receiver.

3. For the $N = \infty$ case, the composite traffic arrival distribution is Poisson with an average arrival rate of ν packets per slot. For the $N < \infty$ case, the composite arrival distribution is derived below by means of a Markovian model.

4. All users employ the same error detecting/correcting code.

5. All users employ a common carrier frequency, code symbol duration T_S, and modulation technique.

6. All users employ the same processing gain $\eta = T_S/T_C$, where T_C is the PN chip duration.

7. Each user has acquired coarse and fine synchronization.

8. The effect of multiple-access interference consists of additional broad-band noise as discussed earlier. Therefore, the equivalent one-sided noise PSD due to j^{th} transmitter at the receiver matched to the desired signal is $N_j = \gamma_j P T_C$, where γ_j is the attenuation of the j^{th} signal and P is the transmitted power. The attenuation can be used to account for fading, or users having unequal transmitted power.

9. The user's signals are mutually noncoherent so that their PSD's add.

10. Because the channel is not a capture channel, all packets arrive at the receiver at the same time and at equal power levels.

Finally, we will represent the total additive noise at the front end of the n^{th} receiver as [9]

$$N_t = N_o + \eta^{-1} \sum_{\substack{j=1 \\ j \neq n}}^{m} E_S \gamma_j, \tag{10}$$

where N_o is the background noise PSD, E_S is the code symbol energy, and m is the number of simultaneous users. Since $E_S = rE_b$, the bit energy-to-total noise ratio for the n^{th} user

is

$$\lambda_t = \frac{\gamma_n E_b}{N_o + \eta^{-1} r E_b \sum_{\substack{j=1 \\ j \neq n}}^{m} \gamma_j} \tag{11}$$

$$= \frac{\gamma_n \lambda_o}{1 + \eta^{-1} r \lambda_o \sum_{\substack{j=1 \\ j \neq n}}^{m} \gamma_j},$$

where λ_o is the bit energy-to-background noise ratio. With the above equal power assumption, (11) becomes

$$\lambda_t = \frac{\lambda_o}{1 + \eta^{-1} r \lambda_o (m-1)}. \tag{12}$$

III Network Analysis

In this section, the steady-state and dynamic performance of the Type I Hybrid ARQ protocol in slotted direct-sequence CDMA systems is analyzed. Expressions for the asymptotic throughput (hereafter referred to as the throughput) and corresponding average packet delay are developed for the steady-state performance. To evaluate the dynamic performance, expressions are developed for the input-output packet flow rates of the network.

III-A Steady-State Throughput-Delay Analysis

The throughput T at the output of the Hybrid ARQ channel is defined as the expected number of successful packets S per slot;

$$T = E\{S\}. \tag{13}$$

By the chain rule of expectation, (13) can be expressed as

$$T = E\{E\{S \mid M\}\}, \tag{14}$$

where

$$E\{S \mid M\} = \sum_{k=1}^{l} k p_{S|M}(k \mid l) \tag{15}$$

and

$$p_{S|M}(k \mid l) = \text{Prob}\{S = k \text{ successful packets} \mid M = l \text{ attempted transmissions}\},$$

$$= \binom{l}{k} P_C^k(l)[1 - P_C(l)]^{l-k}. \tag{16}$$

Here, $P_C(l)$ is the probability of a correctly received packet when there are l simultaneous packet transmissions. Using (15) and (16), the throughput expression in (14) becomes

$$T = \sum_{l=1}^{N} \left\{ \sum_{k=1}^{l} k \binom{l}{k} P_C^k(l)[1 - P_C(l)]^{l-k} \right\} f_M(l), \tag{17}$$

where $f_M(l)$ is the probability distribution function for the number of attempted transmissions (composite arrivals) during a particular slot interval. Since

$$\left\{ \sum_{k=1}^{l} k \binom{l}{k} P_C^k(l)[1 - P_C(l)]^{l-k} \right\} = lP_C(l), \tag{18}$$

then

$$T = \sum_{l=1}^{N} lP_C(l) f_M(l). \tag{19}$$

Among other factors, throughput depends upon the channel's code rate r, processing gain η, and the bit energy-to-background noise ratio λ_o. By combining equations (9) and (12), the following expression can be obtained;

$$m = \frac{\eta}{r} \left(\frac{r}{f^{-1}(P_S)} - \frac{1}{\lambda_o} \right) + 1. \tag{20}$$

where $f^{-1}(\cdot)$ is the inverse operator of $f(\cdot)$ and P_S is the code symbol error probability given by (5) and (7) for the cutoff rate and capacity cases, respectively. As in [3], assume that the code used has the property that the packet error probability $P_E(l) = 1 - P_C(l)$ is zero for $r \leq \{R_o, C\}$ and one if $r > \{R_o, C\}$. For the cutoff rate case, this actually occurs when a very long constraint length convolutional code is used with sequential decoding. Similarly, for the capacity case, this occurs as the coding complexity becomes arbitrarily large. In either case, it is assumed that the coding overhead due to forward error detection (FED) is small and can be neglected. For the limiting case as $r \rightarrow \{R_o, C\}$ and for fixed values of N, $\{R_o, C\}$, η, and λ_o, (20) gives the maximum number of simultaneous packet transmissions as $\hat{m} = \lfloor m \rfloor$. If the actual number of packets l exceeds $\hat{m}$, then information is being transmitted at a rate above the cutoff rate or capacity of the channel. As a result, all packets are incorrectly received and must be retransmitted ($P_C(l) = 0, l > \hat{m}$). If the number of packets is less than or equal to $\hat{m}$, then all packets are received successfully

($P_C(l) = 1, l \leq \hat{m}$). The throughput expression in (19) becomes

$$T = \sum_{l=1}^{\min(\hat{m},N)} l f_M(l). \tag{21}$$

In order to compare throughputs of different systems on an equal basis, we need to account for $\{R_o, C\}$ and η. The normalized throughput or network utilization becomes

$$T(p_r, \{R_o, C\}, \eta) = \frac{\{R_o, C\}}{\eta} T = \frac{\{R_o, C\}}{\eta} \sum_{l=1}^{\min(\hat{m},N)} l f_M(l). \tag{22}$$

This is the average number of successful packets (per slot) per unit time per unit bandwidth. Note that the normalized throughput is a function of the probability of retransmission, the code rate, and the processing gain, and is always less than one. The normalized throughput is also a function of the number of users, the composite traffic intensity, and the bit energy-to-background noise ratio. However, this dependency is not shown explicitly because these parameters are assumed to be uncontrolled. The maximum normalized throughput is

$$\begin{aligned} T_{\{R_o,C\}} &= \max_{0<p_r\leq 1} \max_{\substack{0<\{R_o,C\}\leq 1 \\ 1\leq\eta}} T(p_r, \{R_o, C\}, \eta), \\ &= \max_{0<p_r\leq 1} \max_{\substack{0<\{R_o,C\}\leq 1 \\ 1\leq\eta}} \frac{\{R_o, C\}}{\eta} \sum_{l=1}^{\min(\hat{m},N)} l f_M(l). \end{aligned} \tag{23}$$

For the cutoff rate case, (23) represents the practically achievable upper limit on the maximum normalized throughput T_{R_o}. For the capacity case, (23) represents its theoretically achievable (best) upper limit, T_C. The optimal probability of retransmission, code rate, and processing gain are denoted by p_r^*, $\{R_o^*, C^*\}$, and $\{\eta^{R_o}, \eta^C\}$, respectively.

The composite packet arrival distribution $f_M(l)$ is derived by using a Markov model similar to that developed in [10]. The channel is viewed as a discrete-time system where $X(t)$ represents the number of backlogged users n during a particular slot interval. The network may then be in any one of $N+1$ possible states. A geometric distribution for the probability of retransmission results in the retransmission delay also being geometrically distributed with an average

$$D_r = E/p_r, \tag{24}$$

where E is the average number of packet retransmissions per successfully transmitted packet. The memoryless property of the geometric distribution allows for a simple, single state description. Assuming N and p_o to be time-invariant, $X(t)$ represents a Markov process with state transition matrix $\mathbf{P} = [p_{ij}]$ where the state transition probabilities are defined as

$$p_{ij} = \text{Prob}\{X(t+1) = j \mid X(t) = i\}. \tag{25}$$

Calculation of the above state transition matrix is more complicated than in the narrowband case [10], but is simpler than in [1] because $P_C(l)$ is binary in value ($P_C(l) = 1$ for $l \leq \hat{m}$ and $P_C(l) = 0$ for $l > \hat{m}$). From this simplification, the following expressions for the one-step state transition probabilities can be derived (Appendix A);

$$p_{ij} = \begin{cases} \text{(for } j < i\text{)} & \\ 0, & i - j > \hat{m} \\ \sum_{k=0}^{\min(\hat{m}-i+j,N-i)} b(k, N-i, p_o) b(i-j, i, p_r), & i - j \leq \hat{m} \\ \text{(for } j = i\text{)} & \\ \sum_{k=0}^{\min(N-i,\hat{m})} b(k, N-i, p_o) b(0, i, p_r), & i \leq \hat{m} \\ \sum_{k=0}^{\min(N-i,\hat{m})} b(k, N-i, p_o) b(0, i, p_r) & \\ \quad + \quad \sum_{k=\hat{m}+1}^{i} b(k, i, p_r) b(0, N-i, p_o), & \hat{m} < i < N \\ \sum_{k=\hat{m}+1}^{N} b(k, N, p_r) + b(0, N, p_r), & i = N \\ \text{(for } j > i\text{)} & \\ 0, & j \leq \hat{m}, j - i \leq \hat{m} \\ b(j-i, N-i, p_o)[1 - \sum_{k=0}^{\hat{m}-j+i} b(k, i, p_r)], & \hat{m} < j \leq N, j - i \leq \hat{m} \\ b(j-i, N-i, p_o), & \hat{m} < j \leq N, j - i > \hat{m} \end{cases} \tag{26}$$

where the terms in the summation represent the binomial distribution function such that

$$b(\alpha, n, p) = \binom{n}{\alpha} p^{\alpha} (1-p)^{n-\alpha}. \tag{27}$$

When $\hat{m} = 1$, (26) reduces to the standard narrow-band slotted ALOHA case as in [10]. For the infinite population model, if we let $N \to \infty$ and $p_o \to 0$ such that $Np_o = T_{\{R_o,C\}}$, then (26) becomes

$$p_{ij} = \begin{cases} \text{(for } j < i\text{)} & \\ 0, & i - j > \hat{m} \\ \sum_{k=0}^{\min(\hat{m}-i+j,N-i)} \frac{T^k}{k!} e^{-T} b(i-j,i,p_r), & i - j \le \hat{m} \\ \text{(for } j = i\text{)} & \\ \sum_{k=0}^{\min(N-i,\hat{m})} \frac{T^k}{k!} e^{-T} b(0,i,p_r), & i \le \hat{m} \\ \sum_{k=0}^{\min(N-i,\hat{m})} \frac{T^k}{k!} e^{-T} b(0,i,p_r) & \\ \quad + \quad e^{-T} \sum_{k=\hat{m}+1}^{i} b(0,N-i,p_o), & \hat{m} < i < N \\ \sum_{k=\hat{m}+1}^{N} b(k,N,p_r) \; + \; b(0,N,p_r), & i = N \\ \text{(for } j > i\text{)} & \\ 0, & j \le \hat{m}, j - i \le \hat{m} \\ \frac{T^{j-i}}{(j-i)!} e^{-T} [1 - \sum_{k=0}^{\hat{m}-j+i} b(k,i,p_r)], & \hat{m} < j \le N, j - i \le \hat{m} \\ \frac{T^{j-i}}{(j-i)!} e^{-T}, & \hat{m} < j \le N, j - i > \hat{m} \end{cases} \tag{28}$$

With $\mathbf{P}$, we can solve for the equilibrium state probability distribution $\underline{\pi}$ in the following equation:

$$\underline{\pi} = \underline{\pi}\mathbf{P}, \tag{29}$$

where

$$\underline{\pi} = [\pi(0), \pi(1), \ldots, \pi(N)]. \tag{30}$$

Using the above solution for $\underline{\pi}$, we can now express $f_M(l)$ in terms of the number of backlogged users as

$$f_M(l) = \sum_{n=0}^{N} f_M(l \mid n)\pi(n), \tag{31}$$

where

$$f_M(l \mid n) = \sum_{k=\max(l-n,0)}^{\min(l,N-n)} b(k, N-n, p_o) b(l-k, n, p_r). \tag{32}$$

Substituting (32) into (23), the normalized throughput equation becomes

$$T_{\{R_o,C\}} = \max_{0<p_r\leq 1} \max_{\substack{0<\{R_o,C\}\leq 1 \\ 1\leq\eta}} \frac{\{R_o,C\}}{\eta} \sum_{l=1}^{\min(\hbar,N)} l \sum_{n=0}^{N} \pi(n) \sum_{k=\max(l-n,0)}^{\min(l,N-n)} b(k, N-n, p_o) b(l-k, n, p_r). \tag{33}$$

For the special case when $N = \infty$, $p_o = p_r = 0$ and $p_o N = T_{\{R_o,C\}}$, (33) becomes

$$T_{\{R_o,C\}} = \max_{0<p_r\leq 1} \max_{\substack{0<\{R_o,C\}\leq 1 \\ 1\leq\eta}} \frac{\{R_o,C\}}{\eta} \sum_{l=1}^{\min(\hbar,N)} \frac{l\nu^l \exp\{-\nu\}}{l!} \tag{34}$$

which is the same as the throughput expression derived in [3] where FEC is employed without retransmissions.

The average packet delay D of a network is often modeled as the sum of a random delay component D_r, a deterministic component D_d, and a one slot transmission time:

$$D = D_r + D_d + 1. \tag{35}$$

The deterministic component D_d can be used to model any fixed delay inherent in the network such as the round-trip delay time required by each acknowledgement when using the stop-and-wait ARQ scheme. For this analysis, it is assumed that the deterministic component is zero. As such, equation (35) becomes

$$D = D_r + 1. \tag{36}$$

The random delay component D_r is the average retransmission delay defined earlier and represents the average rescheduling delay or the average backlog time experienced by a user. This random component is necessary to prevent backlogged users from contending with other users blocked during the same time slot. From Little's result [11], D_r can be expressed as

$$D_r = \frac{\bar{n}}{T_{\{R_o,C\}}}, \tag{37}$$

where $\bar{n}$ is the average number of backlogged users which can be calculated from

$$\bar{n} = \sum_{n=0}^{N} n\pi(n). \tag{38}$$

By using (36) and (37), the throughput can be directly related to the average packet delay (35) as

$$T_{\{R_o,C\}} = \frac{\bar{n}}{D-1}. \tag{39}$$

This alternate expression for the throughput is the one that is actually used in the maximization required by (33). The throughput expression in (23) provides a convenient check during the analysis.

Throughput can also be related to E, the average number of packet retransmissions per successfully transmitted packet, by the following expression

$$E = \frac{\nu}{T_{\{R_o,C\}}} - 1, \tag{40}$$

where $\nu/T_{\{R_o,C\}}$ is the average number of times a packet must be transmitted until it is successfully received. This expression can be used with the data from section IV (Fig 4) to obtain a plot of E against $T_{\{R_o,C\}}$. Substituting (40) and (37) into (24), the probability of retransmission can be expressed as

$$p_r = \frac{\nu - T_{\{R_o,C\}}}{\bar{n}}. \tag{41}$$

A similar expression for the probability of original packet transmission can be obtained as

$$\begin{aligned} p_o &= \frac{\nu_o}{\sum_{n=1}^{N}(N-n)\pi(n)} \\ &= \frac{T}{N-\bar{n}} \end{aligned} \tag{42}$$

These expressions for p_r and p_o are used in computing the P matrix (26) and are key in performing the maximization in (23).

III-B Dynamic Throughput-Delay Analysis

The analysis presented so far is not sufficient to fully characterize the performance of our CDMA network. Up to this point, we have assumed that the network is in equilibrium.

That is, for a given number of backlogged users n, the input transmission rate T_{in} is equal to the output or delivered transmission rate T_{out}. The channel input transmission rate is defined as the number of new packet transmissions per time slot. The normalized input transmission rate is

$$T_{in} = \frac{\{R_o, C\}}{\eta}(N - n)p_o. \tag{43}$$

Note that if we take the expectation over n in (43), then T_{in} becomes the input traffic intensity ν_o. The channel output transmission rate is defined as the probability of up to $\hat{m}$ successful packet transmissions in a given time slot. The normalized output transmission rate is given by

$$T_{out} = \frac{\{R_o, C\}}{\eta} \sum_{k=0}^{\hat{m}} \sum_{l=0}^{\hat{m}-k} b(k, N-n, p_o) b(l, n, p_r) - (1-p_o)^{N-n}(1-p_r)^n. \tag{44}$$

Note that if we take the expectation over n in (44), then the output transmission rate becomes the normalized throughput in (22). When N and p_o are stationary with time, (43) and (44) are referred to as the channel load line and equilibrium contour, respectively.

A channel is said to be stable when its load line intersects (nontangentially) the equilibrium contour in exactly one place. Points where the load line intersects the equilibrium contour are defined as channel operating points. For stable channels with a large population size, analysis has shown that networks actually operate close to these channel operating points [12]. In Fig. 3, (44) is plotted against three different channel load lines (43) to show the effects of changing network operating conditions on the network's stability. Load line A represents a stable channel having an operating point (T_1, n_1) with relatively high throughput and a low number of backlogged users (low delay). Load line C also represents a stable channel, but in this case the network is overloaded with a channel operating point having relatively low throughput and a high number of backlogged users (high delay). Load line B represents an unstable channel having three possible channel operating points with the most desirable operating point being at (T_3, n_3). Assuming that the network is currently operating in the vicinity of this point, it will not remain there because $X(t)$ is a random process. That is, there is a nonzero probability that the number of backlogged users will exceed n_2. When this occurs, $T_{in} > T_{out}$ which causes the operating point to drift

(accelerate) up the load line to the overloaded channel operating point (low throughput and high delay). While there is also a nonzero probability that the channel operating point will return below the $\{T_2, n_2\}$ point, simulations show that the channel tends to remain near the overloaded operating point instead [12].

For the finite population case, if the channel is stable, then the network operating point can reside in the vicinity of the channel operating point indefinitely. For an unstable channel, however, network operation around the desired channel operating point is achievable for only a finite amount of time before the channel becomes overloaded. Finite population channels can be stabilized by decreasing p_r (at the expense of increasing delay), by decreasing N, or by employing external stabilizing measures which include various retransmission control strategies [13]. For the infinite population case, channels are inherently unstable and can operate near their desirable operating points for only a finite amount of time [10,14]. This operational time can be increased by the methods mentioned above to achieve performance close to the theoretical optimum [13].

IV Performance Evaluation

In this section, the dependency of the network's throughput-delay performance on N, ν, λ_o, p_r, $\{R_o, C\}$, and η is examined in detail for both the steady-state and dynamic cases. The optimal probabilities of retransmission, processing gains, and code rates are identified. The results obtained here for DS/DPSK signaling can also be obtained for DS/BPSK signaling within the same framework.

IV-A Steady-State Throughput-Delay Performance

Figs. 4 and 5-6 are plots of the maximum normalized throughtputs, T_{R_o} and T_C, against traffic intensity ν and against average delay D, respectively. These plots illustrate the combined effects of maximizing throughput over the probability of retransmission p_r, and the code rate and processing gain $\{R_o, C; \eta\}$. The first maximization required by (23) is performed by fixing D in (39), allowing p_r to vary ($0 \leq p_r \leq 1$) in (29), and solving for a $(\bar{n}, \pi(n))$ solution which maximizes (39). This procedure is repeated for all possible values

of $\hat{m}$ ($1 \leq \hat{m} \leq N$) and results in N optimal $(p_r, T_{\{R_o,C\}})$ pairs for each fixed value of D. An alternate procedure, whereby the throughput is fixed and the delay is minimized, produces the same optimal $(p_r, T_{R_o,C}, \hat{m}, D)$ combinations. For the second maximization required by (33), D is again fixed for a given level of background noise. The code rate and processing gain $\{R_o, C; \eta\}$ are allowed to vary in (20) yielding $\hat{m}$ values for which the corresponding throughputs are known from step one. Each value of throughput is then normalized and the maximum normalized value is selected. Note that (20) requires a solution for P_S. For the cutoff rate case, a solution for P_S is found in a straight forward manner by using (5) and (6). For the capacity case, a P_S solution is found by solving (3) and (4) simultaneously.

Results in Figs. 4 and 5-6 are shown for population sizes $N = 10, 30$, and ∞; and for background noise levels of $\lambda_o = \infty$ and $\lambda_o = 10.34$ dB. The background noise level of 10.34 dB was chosen as it results in a bit error probability of 10^{-5} for uncoded DPSK in the absence of multiple-access interference. Although the infinite population case may be an unrealistic population size, it does provide a useful bound on system performance.

IV-A.1 Optimization of Throughput-Delay Performance over Probability of Retransmission

Optimal values for probability of retransmission p_r^* exist for all values of traffic intensity as shown in Fig. 7. For the low range of traffic intensities ($\nu < 1.0$), there are two distinct p_r^* values for each value of ν. Selecting the lower p_r^* value for a given traffic intensity results in the maximum possible value for $\{T_{R_o}, T_C\}$ (Fig. 4) but with nonminimal delay (Figs. 5-6). In Figs. 5-6, these lower p_r^* values correspond to the upper branch of the throughput-delay curve. Here, ν is increasing along the curve from A to B and back to C. Points on the A-B portion of the curve correspond to stable operating points, whereas those on the B-C portion of the curve do not. A maximal $\{T_{R_o}, T_C\}$ value is achieved under the following conditions: $p_r^* \ll 1.0$, $p_o \rightarrow 1.0$, and $\bar{n} \rightarrow N$. As a result, the throughput is closely approximated by the equation $\{T_{R_o}, T_C\} = \nu$ for low ν ($\nu < 0.5$) in Fig. 4, and by the equation $\{T_{R_o}, T_C\} = N/D$ for the upper branch (A-B) in Figs. 5-6. These two throughput expressions are related through Little's result (37) by applying the above conditions to (41)

and (39), respectively.

Selecting the higher p_r^* value for a given traffic intensity ($\nu < 1.0$) in Fig. 7 results in a nonmaximal $\{T_{R_o}, T_C\}$ value (not shown in Fig. 4) but at minimal delay (Figs. 5-6). The higher p_r^* values correspond to the lower right-hand branch of the throughput-delay curve in Figs. 5-6 (increasing ν from E to F). This throughput-delay curve agrees with the results from earlier throughput-delay analyses for conventional narrow-band slotted ALOHA [12]. For these nonmaximal $\{T_{R_o}, T_C\}$ values, $p_r^* > p_o$ and $\bar{n}$ is relatively low.

Note in Fig. 7 that for the low range of traffic intensities ($\nu < 1.0$), the upper and lower curves for p_r^* are separate and distinct. For a given traffic intensity, this separation represents the change in p_r^* that is necessary to move the network from an operating point having maximal throughput to one having minimal delay, and vice-versa. In Fig. 5, for example, the throughput-delay points for traffic intensities of $\nu = 0.525$ and 1.00 are plotted for comparison purposes for the two values of p_r^*. Note in this example, that for $\nu = 1.0$ it seems that an improvement in throughput can be had for a modest increase in delay by operating on the upper branch (B-C). However, the upper branch (B-C) operating point is not stable, whereas, the lower branch (E-F) operating point is. Stability issues are discussed further in section IV-B. Note that for this range of ν, p_r^* values do not depend on λ_o.

For the upper range of traffic intensities ($\nu > 1.0$) in Fig. 7, there are also two possible optimal values for the probability of retransmission p_r^* as indicated by the solid and dashed lines. The solid line corresponds to the case when either plain or Hybrid ARQ is the optimal protocol. Here, p_r^* can be approximated by $p_r^* = \nu/N$. Because p_r^* is independent of λ_o, the same value of p_r^* is used for both plain and Hybrid ARQ for a given traffic intensity. The dashed line corresponds to the case when Type 1 Hybrid ARQ with CDMA (hereafter referred to as CDMA) is used. Here, $p_r^* = \nu/N$ exactly and the composite arrival distribution $f_M(l)$ is binomial. Note in Fig. 7 that the (dashed) $p_r^* = \nu/N$ lines are shown for their widest possible range of values. Where they actually begin (lower left end) depends on the cutoff intensity (discussed below) which, in turn, depends on λ_o. Where the dashed lines end (upper right end) depends on the population size and the traffic intensity. At high intensity, high processing gain causes $\hat{m} = N$. As a result, there are no packets to

retransmit and the $p_r^* = \nu/N$ lines terminate.

In Fig. 4, the plain/Hybrid ARQ p_r^* values correspond to the plain/Hybrid ARQ curves and the CDMA p_r^* values correspond to the upper branches which break away from the plain ARQ curves at various traffic intensities. In Figs. 5-6, the B-C-A branch of the throughput-delay curve (ν increasing from B to C then toward A) corresponds to the plain/Hybrid ARQ p_r^* values and the C-D branch corresponds to the CDMA p_r^* values.

Note in Figs. 4 and 5-6 that for the plain/Hybrid ARQ case, throughput becomes vanishingly small and delay becomes unbounded as the traffic intensity increases. However, when CDMA is used, this degradation in performance does not occur. Instead, the maximum normalized throughput (Fig. 4) decreases to some minimal value, at a particular traffic intensity, and then increases again. The corresponding delay decreases (Figs. 5-6). The traffic intensity at which CDMA achieves this improved performance over plain/Hybrid ARQ is defined as the *cutoff traffic intensity* and is denoted as $\{\nu^{R_o}, \nu^C\}$. In Fig. 4, the cutoff intensities occur at the intersection of the CDMA curves with the plain/Hybrid ARQ curves. Note that for $\lambda_o = \infty$ and $\nu < \{\nu^{R_o}, \nu^C\}$, T_{R_o} and T_C are identically equal. For $\lambda_o < \infty$ and $\nu < \{\nu^{R_o}, \nu^C\}$, $T_{R_o} \neq T_C$, but these differences are usually too small to show graphically, the only exception being for T_{R_o} at $\lambda_o = 10.34$ dB. In Figs. 5-6, point C is an example of a throughput-delay point which corresponds to a cutoff intensity. Note how the delay drops precipitously (C-D) when CDMA is used. Recall that the above throughput-delay results must be adjusted for any fixed delay inherent in the particular network being considered. Additional fixed delay causes the curves in Figs. 5-6 to be shifted upward by a corresponding amount.

IV-A.2 Optimization of Throughput-Delay Performance over Code Rate and Processing Gain

An important parameter arising from the maximization of the normalized throughput over the processing gain and code rate in (33) is the cutoff traffic intensity, $\{\nu^{R_o}, \nu^C\}$, described as follows. If $\nu < \{\nu^{R_o}, \nu^C\}$, then the optimal processing gain $\{\eta^{R_o}, \eta^C\}$ is unity for all bit energy-to-background noise ratios ($\lambda_o \leq \infty$). If $\nu > \{\nu^{R_o}, \nu^C\}$, then $\{\eta^{R_o}, \eta^C\} > 1$ with

the actual value depending on the traffic intensity, population size, and the bit energy-to-background noise ratio as described below.

Fig. 8 illustrates the effect of bit energy-to-background noise ratio λ_o and population size N on the cutoff traffic intensity $\{\nu^{R_o}, \nu^C\}$. In general, $\{\nu^{R_o}, \nu^C\}$ increases to its maximum value (N) as λ_o decreases. Note that for each $\{\nu^{R_o}, \nu^C\}$ there is a corresponding λ_o which is defined as the *noise limit* for that cutoff intensity, denoted by $\{\lambda_o^{R_o}, \lambda_o^C\}$. Also, for each λ_o there is a corresponding cutoff traffic intensity. The {6.82 dB, 4.81 dB} noise limits are defined as the network *asymptotic noise limits* and are denoted $\{\hat{\lambda}_o^{R_o}, \hat{\lambda}_o^C\}$. The network cannot operate at bit energy-to-background noise ratios below the asymptotic noise limits, because they represent the smallest λ_o that can be present for reliable coded communications in the absence of multiple-access interference [3].

Fig. 9 shows how the optimal processing gain $\{\eta^{R_o}, \eta^C\}$ depends on the bit energy-to-background noise ratio λ_o for different traffic intensities ν and population sizes N. In general, the optimal processing gain increases with decreasing λ_o. For the $N = \infty$ case, $\{\eta^{R_o}, \eta^C\}$ increases toward infinity as λ_o decreases to the asymptotic noise limit $\{\hat{\lambda}_o^{R_o}, \hat{\lambda}_o^C\}$. For the $N < \infty$ case, decreasing λ_o causes a similar increase in $\{\eta^{R_o}, \eta^C\}$, but toward a finite maximum value at the noise limit corresponding to a given traffic intensity. For a particular traffic intensity, if $\{\hat{\lambda}_o^{R_o}, \hat{\lambda}_o^C\} < \lambda_o < \{\lambda_o^{R_o}, \lambda_o^C\}$ then the optimal processing gain is unity. Note that for most values of λ_o, $\{\eta^{R_o}, \eta^C\}$ depends primarily on ν and not so much on the population size.

Figs. 10-11 show how the optimal processing gain $\{\eta^{R_o}, \eta^C\}$ depends on the traffic intensity for the different population sizes N and bit energy-to-background noise ratios λ_o. For the $N = \infty$ case, $\{\eta^{R_o}, \eta^C\}$ depends on ν in a nearly linear fashion. The effect of finite population size is a 'staircase' type curve which tends to the $N = \infty$ curve as $N \to \infty$.

Fig. 12 summarizes the dependency of the optimal code rate $\{R_o^*, C^*\}$ on the bit energy-to-background noise ratio λ_o and population size N. Note that the selection of the optimal code rate depends primarily upon ν and λ_o, and does not directly depend on N. When $\lambda_o = \infty$ and $\nu < \{\nu^{R_o}, \nu^C\}$, the optimal code rate $\{R_o^*, C^*\}$ is unity. Hence, the optimal protocol is plain ARQ. In fact, plain ARQ is used *only* when $\lambda_o = \infty$ and

$\nu < \{\nu^{R_o}, \nu^C\}$. When $\lambda_o < \infty$ and $\nu < \{\nu^{R_o}, \nu^C\}$, $\{0.48, 0.50\} < \{R_o^*, C^*\} < 1.0$ In this case, Hybrid ARQ is used. For $\lambda_o \leq \infty$ and $\nu > \{\nu^{R_o}, \nu^C\}$, the optimal code rate $\{R_o^*, C^*\}$ is $\{0.48, 0.50\}$, respectively. CDMA is the optimal protocol in this case.

It should be noted that for the entire range of traffic intensities, it is possible to operate CDMA with a delay $D = 1.0$. Unity delay is achieved by selecting the appropriate code rate and processing gain combination so that $\hat{m} = N$. In most cases, however, unit delay is achieved at the expense of a far from optimal throughput. In general, the throughput at unit delay $T_{D=1}$ is very small for $0 < \nu < \{\nu^{R_o}, \nu^C\}$. At $\nu = \{\nu^{R_o}, \nu^C\}$, there is a step increase in $T_{D=1}$, after which $T_{D=1} \rightarrow T_{\{R_o, C\}}$ as $\nu \rightarrow N$.

Finally, note that for a given level of background noise, the optimal code rate can be determined as the traffic intensity varies by using Figs. 8 and 12 . For a given λ_o, Fig. 8 gives the corresponding $\{\nu^{R_o}, \nu^C\}$. As the traffic intensity varies above and below this particular $\{\nu^{R_o}, \nu^C\}$, the optimal code rate is given by the lower branch and upper branch of Fig. 12, respectively.

IV-B Dynamic Throughput-Delay Performance

When the traffic intensity of our network rises to a level at or above the cutoff intensity, the employment of CDMA stabilizes what would otherwise be an unstable network. This effect is shown in Fig. 13 for a traffic intensity $\nu = 3.0$, a population size of $N = 10$, and no background noise ($\lambda_o = \infty$). The cutoff intensity in this case is $\nu^{R_o} = 2.85$. Note that the plain ARQ curves display an unstable characteristic with an overloaded channel operating point K, whereas the CDMA curves display a stable characteristic with a singular channel operating point J. Average operating points generated by numerical analysis are shown for each case. Points G ($T_{R_o} = 0.129, \bar{n} = 7.417$) and H ($T_{R_o} = 0.151, \bar{n} = 2.70$) correspond to the plain ARQ and CDMA cases, respectively. For the CDMA case, note that the relatively large distance between the average operating point H and the channel operating point J can be attributed to the small population size. In Fig. 5, throughput-delay points corresponding to $\nu = 3.0$ and $N = 10$ for both the plain ARQ ($T_{R_o} = 0.129, D = 56.0$) and CDMA ($T_{R_o} = 0.151, D = 19.0$) cases are shown for reference. In Fig. 6, a throughput-delay

point for $\nu = 3.0$ and $N = 30$ is also shown. Observe that for the $(T_{R_o}, \lambda_o = \infty)$ curve, there is no $\nu = 3.0$ throughput-delay point on the CDMA curve (C-D) because $\nu = 3.0$ is less than the cutoff intensity for $N = 30$, $\lambda_o = \infty$ ($\nu^{R_o} = 3.17$). Thus, for a given traffic intensity and λ_o, an increase in N causes a corresponding increase in $\{\nu^{R_o}, \nu^C\}$ which, in this case, destabilizes the network.

Finally, note that CDMA achieves a similar stabilizing effect for $\lambda_o < \infty$ and $\nu > \{\nu^{R_o}, \nu^C\}$. For this case, Hybrid ARQ operates at some code rate less than unity depending on the background noise level. In the above example ($N = 10$, $\nu = 3.0$), if λ_o is high enough so that $R_o \approx 1.0$ then R_o is close to unity and η is unity in (43) and (44) and the resulting Hybrid ARQ stability curves are very close to those shown for the plain ARQ case in Fig. 13. That is , the Hybrid ARQ case exhibits the same unstable characteristic (ie., multiple network operating points). If 15.25dB $\leq \lambda_o < \infty$ then $R_o < 1.0$ and $\eta = 1.0$ and the resulting equilibrium curves generally shift to the left but maintain the same unstable form. For the CDMA case, 15.25dB $\leq \lambda_o < \infty$ causes the CDMA stability curves to shift in a similar manner while maintaining the same stable form.

V Operational Considerations

The results presented above suggest that the network will operate in one of three modes: plain ARQ, Hybrid ARQ, or CDMA. The mode of operation selected depends primarily upon the traffic intensity ν, the level of background noise λ_o, and, to a lesser degree, the population size N. All of these quantities must be monitored by each user for proper network operation. For the case when the traffic intensity is low enough to permit stable network operation (eg., $\nu < 0.5$), the network will operate in the plain ARQ mode when $\lambda_o = \infty$ and in Hybrid ARQ mode when $\lambda_o < \infty$. In a practical system with finite packet lengths, the plain ARQ mode could be used for a sufficiently high λ_o. In this case, a user will have to decide whether it wants to operate with optimal delay (upper p_r^* values in Fig. 7) or optimal throughput (lower p_r^* values in Fig. 7) by choosing the appropriate probability of retransmission. While in either mode of operation, the probability of retransmission is adjusted according to variations in the traffic intensity and the population size (Fig. 7) and

the code rate is adjusted according to variations in the background noise level (Fig. 12).

When the traffic intensity reaches a certain level ($0.5 \leq \nu \leq 1.0$), the network becomes unstable. As the traffic intensity continues to increase, a cutoff intensity is eventually reached where the optimal protocol becomes CDMA. In order to determine when to switch to CDMA, an upper traffic intensity threshold is set at some value between $\nu \approx 0.5$ and $\nu = \{\nu^{R_o}, \nu^C\}$. This threshold point will most likely be set closer to $\nu \approx 0.5$ than $\nu = \{\nu^{R_o}, \nu^C\}$ to allow the system time to change over to the CDMA mode of operation. Note that a lower threshold traffic intensity must also be set within the $\nu \leq 0.5$ traffic intensity range to trigger the return to the plain/Hybrid ARQ mode. During the CDMA mode of operation, the probability of retransmission is adjusted according to variations in the traffic intensity and population size (Fig. 7), and the processing gain is adjusted according to changes in the traffic intensity, the background noise level, and the population size (Figs. 8-11). The code rate remains constant (Fig. 12).

Finally, it should be noted that the code rate and processing gain will be chosen so that the communication link will be operating at the cutoff rate (capacity) if and only if the actual number of packets in a slot, l, is equal to $\hat{m}$. If $\hat{m}$ is relatively large, then $\text{Prob}(l = \hat{m})$ may be small so that the system operates either above or below the cutoff rate (capacity) most of the time. Hence, the asymptotic performance bounds, especially those in terms of the cutoff rate, may be 'more achievevable' in CDMA systems than in narrow-band systems where $\hat{m} = 1$.

VI Concluding Remarks

This paper has presented an analysis of the steady-state and dynamic performance of Hybrid ARQ protocols in slotted direct-sequence CDMA channels. Numerical analyses were presented to show how network design parameters such as the probability of retransmission, code rate, and processing gain should be chosen in order to maximize system performance. It was also shown how the above design parameters depend on the uncontrolled network parameters such as population size, traffic intensity, and bit energy-to-background noise ratio. Important parameters such as the cutoff intensities, noise limits and asymptotic noise

limits arising from this analysis helped illustrate when it is best to employ one of the various error control techniques discussed in this paper (ARQ, Hybrid ARQ , or CDMA). Many important properties were revealed that are, quite likely, general characteristics of CDMA networks. These include the general nature of the throughput versus traffic intensity and throughput versus delay characteristics, the dependency of the probability of retransmission on the traffic intensity and the critical role that the probability of retransmission plays in the throughput-delay-stability tradeoff, the linear dependency of the optimal processing gain on the traffic intensity, the dependency of the optimal code rate on the bit energy-to-background noise ratio, the relative insensitivity of the system's performance to changes in the number of transceivers, and the stabilizing effects of employing CDMA.

A Appendix

In this appendix, the state transition probabilities p_{ij} are derived for the Markov model discussed in section III-A. Recall that we view the channel as a discrete-time system where $X(t)$ represents the number of backlogged users n during a particular slot interval. For a population size of N paired transceivers, the network may be in any one of $N+1$ possible states. The state transition probability p_{ij} is defined as the probability of moving from a state having i backlogged users in a given time slot to a state having j backlogged users in the next time slot or

$$p_{ij} = \text{Prob}\{X(t+1) = j \mid X(t) = i\}. \tag{45}$$

In one step, $X(t)$ may decrease by an amount up to $\hat{m}$, remain the same, or increase in size up to N.

Case I : A decrease in backlog.

An $l = i - j$ decrease in $X(t)$ occurs when l retransmissions are successful. In the present system, this can occur only when l retransmissions occur together with up to $\hat{m} - l$ new transmissions. Any combination of new transmissions greater than $(\hat{m} - l)$ plus l retransmissions would be blocked by the channel. Any combinations involving fewer or greater than l retransmissions would represent a different state transition probability. Thus,

$$\begin{aligned} p_{i,i-l} &= \text{Prob[up to } (\hat{m} - l) \text{ NTX]} \cdot \text{Prob[}l \text{ RTX]} \\ &= \sum_{k=0}^{\min(\hat{m}-l,N-i)} b(k, N-i, p_o) b(l, i, p_r), \end{aligned} \tag{46}$$

where the terms in the summation represent the binomial distribution function such that

$$b(\alpha, n, p) = \binom{n}{\alpha} p^{\alpha} (1-p)^{n-\alpha}, \tag{47}$$

and NTX and RTX are abbreviations for new packet transmissions and packet retransmissions, respectively. Note that $\hat{m}$ limits the size in which decreases in the number of backlogged users can be made. That is

$$p_{i,i-l} = 0, \qquad \text{for } l > \hat{m} \tag{48}$$

Case II : Backlog remains constant.

Remaining at the same number of backlogged users can occur in three ways. Transmissions may involve neither NTXs nor RTXs, only NTX, or only RTXs. Note that transmissions involving both NTXs and RTXs necessarily increase or decrease $X(t)$.

A. A transmission involving neither NTXs nor RTXs occurs with probability

$$\begin{aligned} p_{ii} &= \text{Prob[no NTX]} \cdot \text{Prob[no RTX]} \\ &= b(0, N-i, p_o) b(0, i, p_r). \end{aligned} \tag{49}$$

B. A transmission involving only NTXs can maintain the same number of backlogged users only if the number of NTXs are less than or equal to $\hat{m}$. This occurs with probability

$$\begin{aligned} p_{ii} &= \text{Prob[no RTX]} \cdot \text{Prob[up to } \hat{m} \text{ NTX]} \\ &= (1-p_r)^i \sum_{k=0}^{\min(N-i,\hat{m})} b(k, N-i, p_o). \end{aligned} \tag{50}$$

C. A transmission involving only RTXs can maintain a constant backlog size if more than $\hat{m}$ RTXs are attempted. This occurs with probability

$$\begin{aligned} p_{ii} &= \text{Prob[no NTX]} \cdot \text{Prob[more than } \hat{m} \text{ RTX]} \\ &= b(0, N-i, p_o) \sum_{k=\hat{m}+1}^{i} b(k, i, p-r). \end{aligned} \tag{51}$$

Using (49), (50), and (51) and accounting for the relative sizes of the current number of backlogged users i, $\hat{m}$, and N, p_{ii} can be expressed as

$$p_{ii} = \begin{cases} \sum_{k=0}^{\min(N-i,\hat{m})} b(k, N-i, p_o) b(0, i, p_r), & i \le \hat{m} \\ \sum_{k=0}^{\min(N-i,\hat{m})} b(k, N-i, p_o) b(0, i, p_r) & \\ \quad + \sum_{k=\hat{m}+1}^{i} b(k, i, p_r) b(0, N-i, p_o), & \hat{m} < i < N \\ \sum_{k=\hat{m}+1}^{N} b(k, N, p_r) + b(0, N, p_r), & i = N \end{cases} \tag{52}$$

Case III : An increase in backlog.

An l increase in the number of backlogged users occurs when l NTXs occur together with more than $(\hat{m}-l)$ RTXs. This occurs with probability

$$\begin{aligned} p_{ii} &= \text{Prob}[l \text{ NTX}] \cdot \text{Prob}[\text{more than } (\hat{m}-l) \text{ RTX}] \\ &= b(l, N-i, p_o)[1 - \sum_{k=0}^{\hat{m}-l} b(k, i, p_r)]. \end{aligned} \tag{53}$$

Note that as opposed to the decrease in backlog case, a one-step increase in backlog is limited by the quantity $N-i$. Note also that increases in backlog are impossible for combinations of NTXs and RTXs totaling less than or equal to $\hat{m}$.

Combining (46), (52), and (53), the state transition probabilities p_{ij} become:

$$p_{ij} = \begin{cases} \text{(for } j < i\text{)} & \\ 0, & i-j > \hat{m} \\ \sum_{k=0}^{\min(\hat{m}-i+j, N-i)} b(k, N-i, p_o) b(i-j, i, p_r), & i-j \le \hat{m} \\ \text{(for } j = i\text{)} & \\ \sum_{k=0}^{\min(N-i, \hat{m})} b(k, N-i, p_o) b(0, i, p_r), & i \le \hat{m} \\ \sum_{k=0}^{\min(N-i, \hat{m})} b(k, N-i, p_o) b(0, i, p_r) & \\ \quad + \sum_{k=\hat{m}+1}^{i} b(k, i, p_r) b(0, N-i, p_o), & \hat{m} < i < N \\ \sum_{k=\hat{m}+1}^{N} b(k, N, p_r) + b(0, N, p_r), & i = N \\ \text{(for } j > i\text{)} & \\ 0, & j \le \hat{m}, j-i \le \hat{m} \\ b(j-i, N-i, p_o)[1 - \sum_{k=0}^{\hat{m}-j+i} b(k, i, p_r)], & \hat{m} < j \le N, j-i \le \hat{m} \\ b(j-i, N-i, p_o), & \hat{m} < j \le N, j-i > \hat{m} \end{cases} \tag{54}$$

Note that for $\hat{m} = 1$, (54) reduces to the state transition probabilities for the conventional narrow-band slotted ALOHA [10], and that (54) is a special case of the state transition probabilities derived in [1] for random access CDMA networks.

References

[1] D. Raychaudhuri, "Performance Analysis of Random Access Packet-Switched Code Division Multiple Access Systems," *IEEE Transactions on Communications*, vol. COM-29, pp. 895–901, June 1981.

[2] A. Polydoros and J. Silvester, "Slotted Random Access Spread-Spectrum Networks: An Analytical Framework," *IEEE Journal on Selected Areas in Communications*, vol. SAC-5, pp. 989–1001, July 1987.

[3] G. L. Stuber, "Asymptotic Throughput in Slotted Direct-Sequence Code Division Multiple-Access Networks," *to be published*, 1988.

[4] G. L. Stuber and J. M. Hanratty, "Asymptotic Performance of Hybrid ARQ Protocols in Slotted Direct-Sequence Code Division Multiple-Access Networks: Jamming Analysis," *to be published*, 1989.

[5] R. A. Comroe and D. J. Costello, "ARQ Schemes for Data Transmission in Mobile Radio Systems," *IEEE Transactions on Vehicular Technology*, vol. VT-33, pp. 88–97, August 1984.

[6] D. C. et al., "Application of code combining to a selective-repeat ARQ link," *MILCOM Conference Record*, vol. 1, pp. 247–252, August 1985.

[7] D. Chase, "Code Combining - A Maximum Likelihood Decoding Approach for Combining an Arbitrary Number of Noisy Packets," *IEEE Transactions on Communications*, vol. COM-33, pp. 385–393, May 1985.

[8] M. B. Pursley, "Performance evaluation for phase-coded spread spectrum multiple-access - Part I: System analysis," *IEEE Transactions on Communications*, vol. COM-25, pp. 795–799, August 1977.

[9] C. L. W. et al., "Performance Considerations of Code Division Multiple-Access Systems," *IEEE Transactions on Vehicular Technology*, vol. VT-30, pp. 3–9, February 1981.

[10] L. Klienrock and S. S. Lam, "Packet Switching in a Multiaccess Broadcast Channel: Performance Evaluation," *IEEE Transactions on Communications*, vol. COM-23, pp. 410–421, April 1975.

[11] L. Klienrock, *Queueing Systems*. New York: John Wiley & Sons, Inc., 1975.

[12] S. S. Lam, *Packet Switching in a Multi-access Broadcast Channel with Application to Satellite Communication in a Computer Network*. PhD thesis, Dep. Comput. Sci., Univ. Calif., Los Angeles, 1974.

[13] S. S. Lam and L. Klienrock, "Packet switching in a multiaccess broadcast channel: dynamic control procedures," *IEEE Transactions on Communications*, vol. COM-23, pp. 891–905, September 1975.

[14] M. Kaplan, "A sufficient condition for nonergodicity of a Markov chain," *IEEE Transactions on Information Theory*, vol. IT-25, pp. 470–471, July 1979.

List of Illustrations

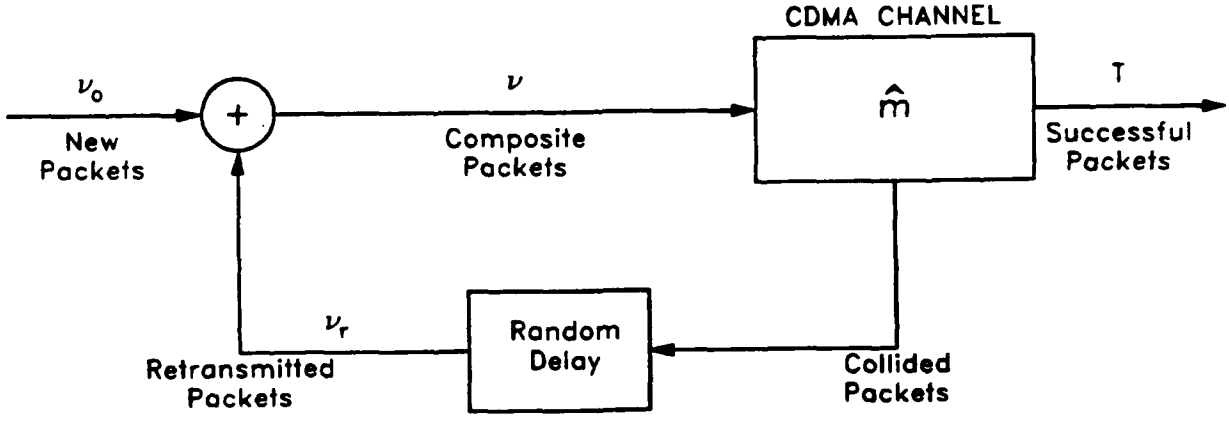


Figure 1: CDMA Network Packet Flow Model

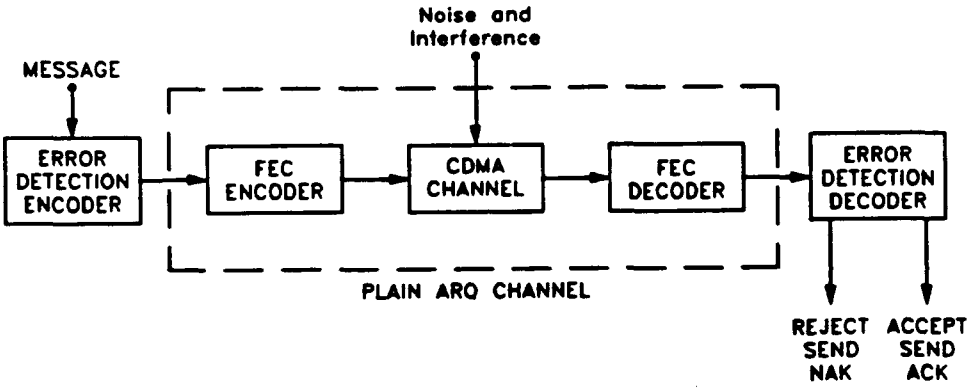


Figure 2: Type I Hybrid ARQ Channel

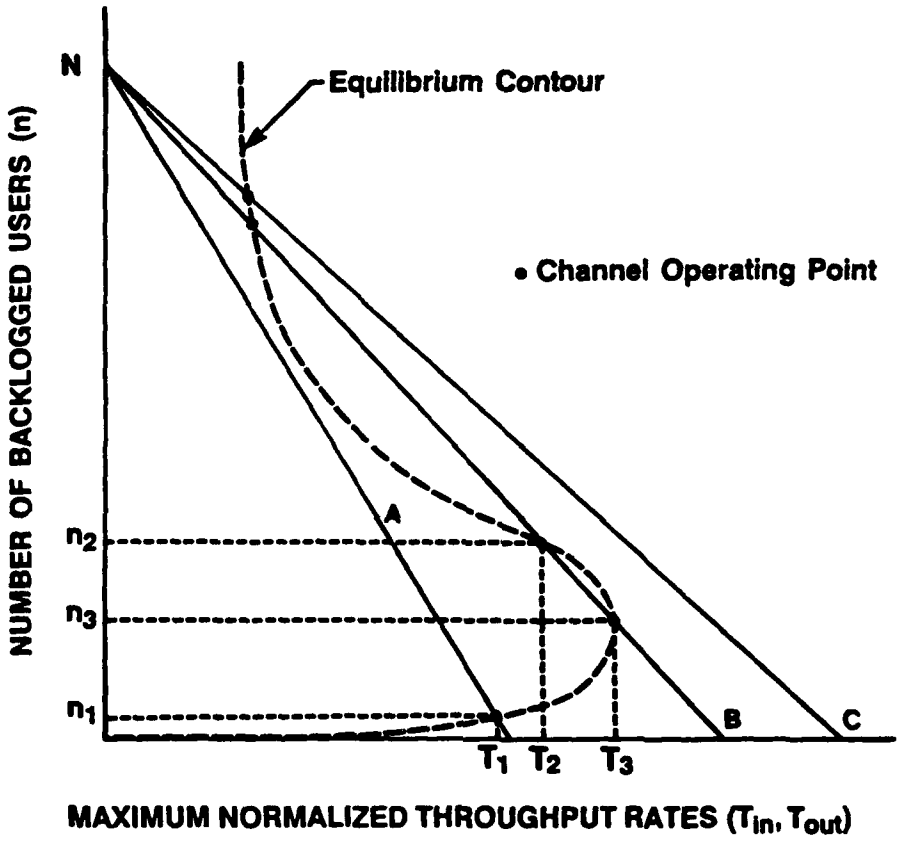


Figure 3: CDMA Network Stability for Different Network Operating Conditions

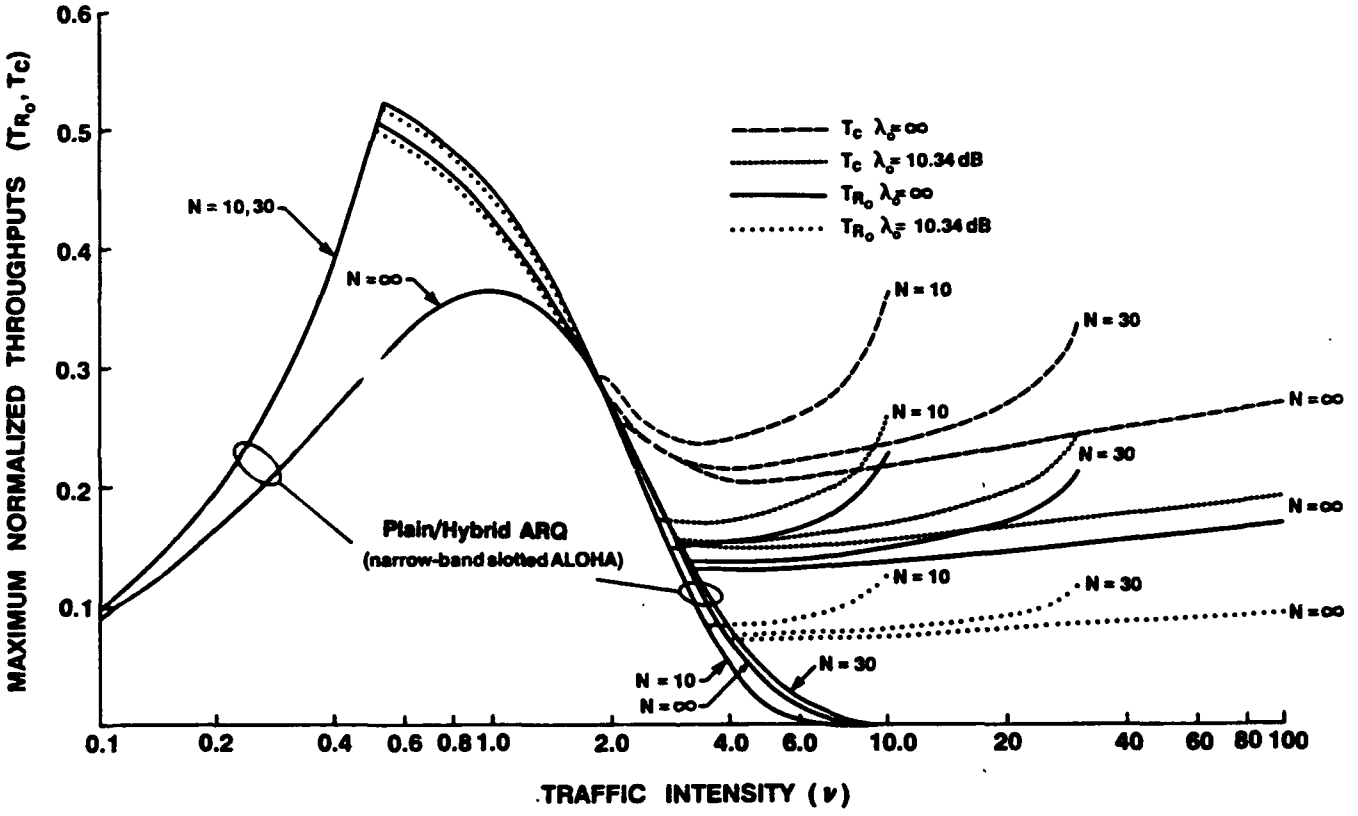


Figure 4: Maximum Normalized Throughput versus Traffic Intensity

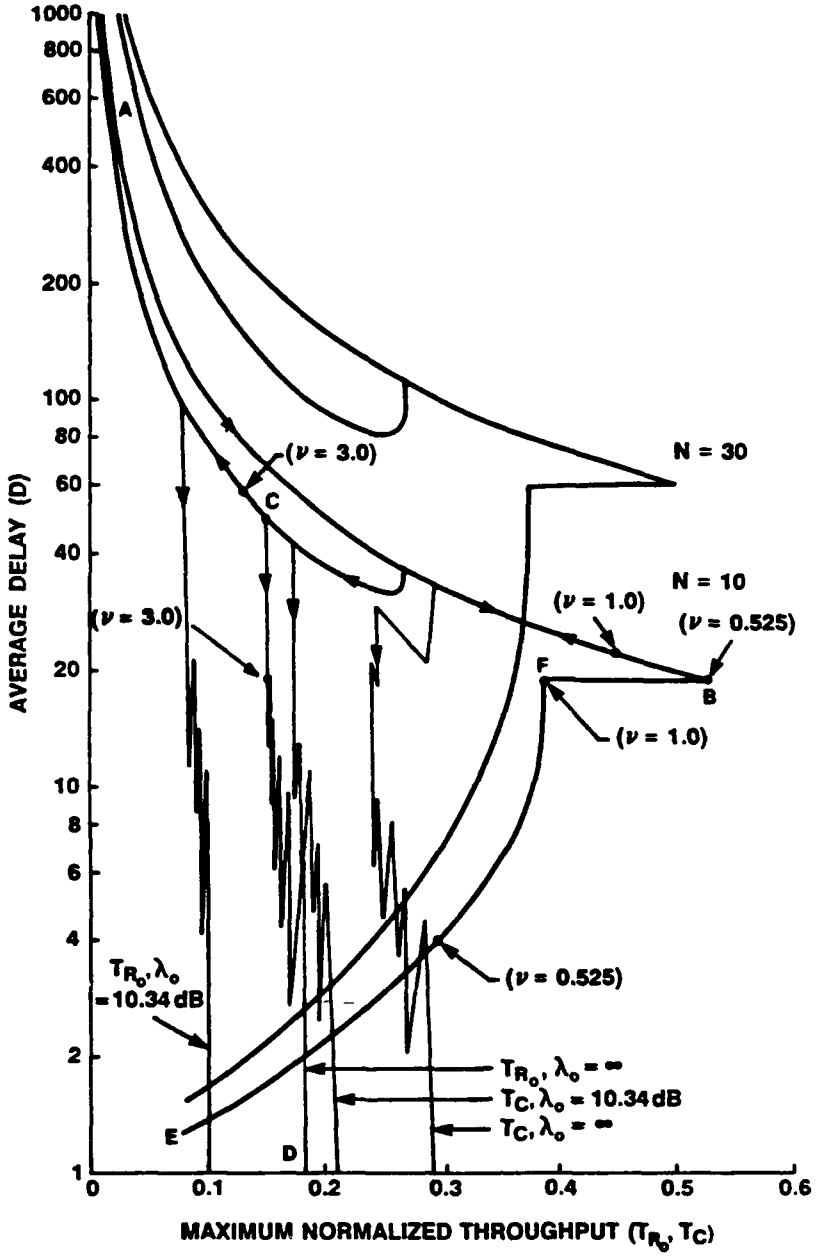


Figure 5: Average Delay versus Maximum Normalized Throughput for N = 10 (N = 30 results are overlayed)

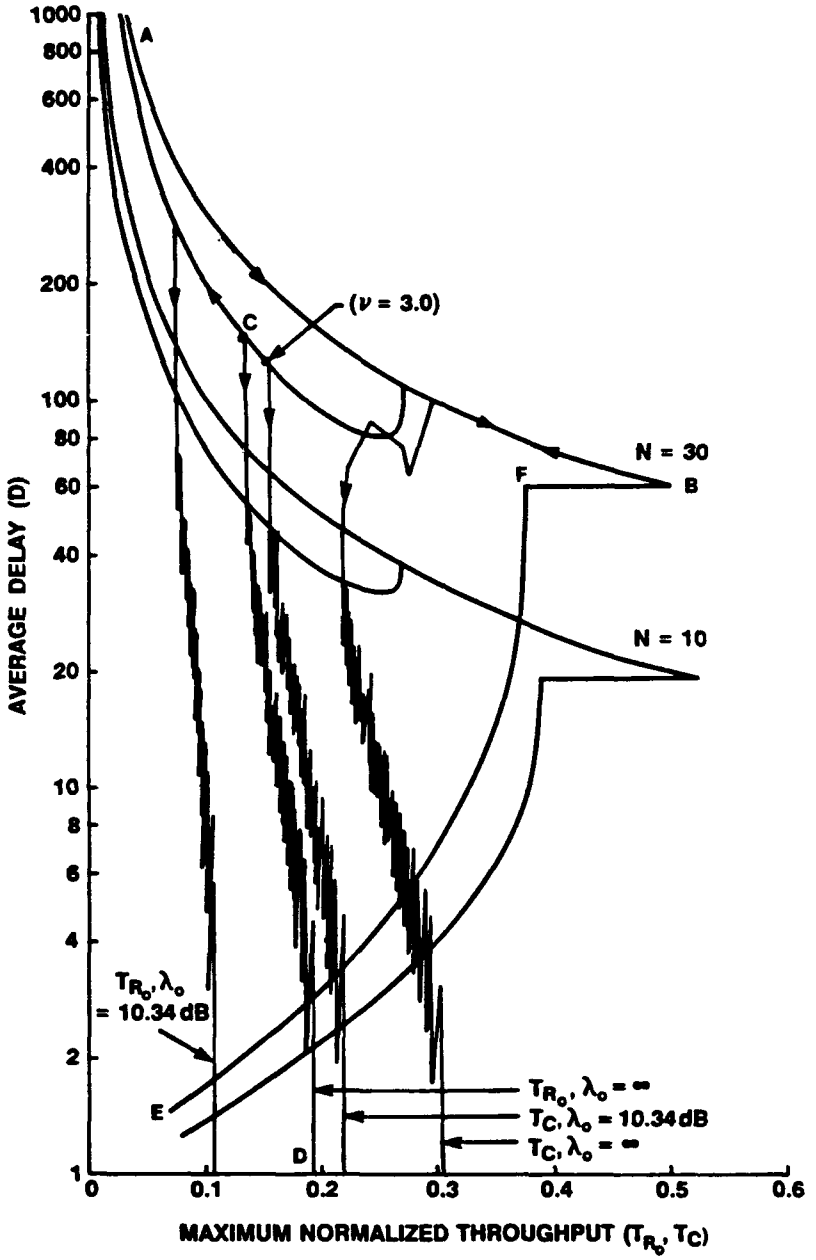


Figure 6: Average Delay versus Maximum Normalized Throughput for N = 30
(N = 10 results are overlayed)

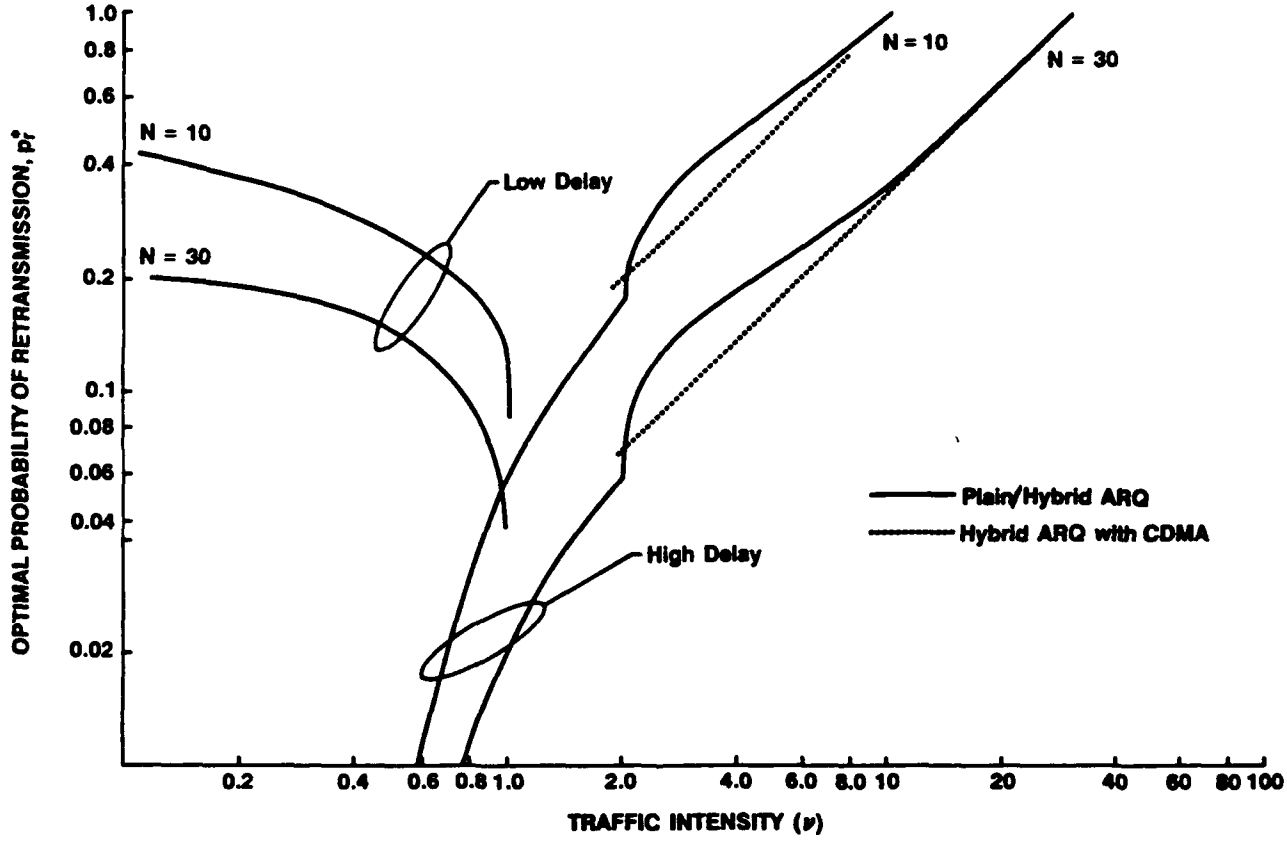


Figure 7: Optimal Probability of Retransmission versus Traffic Intensity

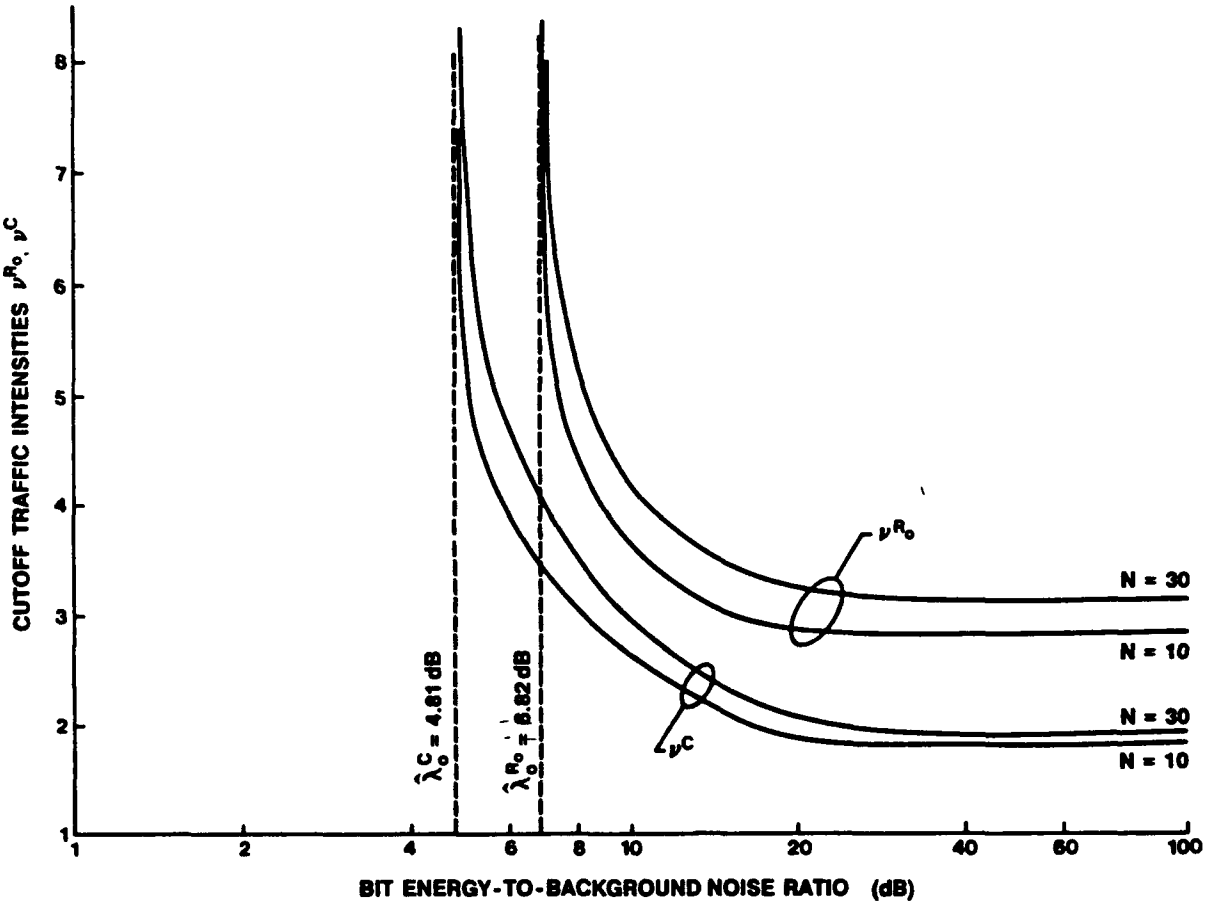


Figure 8: Cutoff Traffic Intensity versus Bit Energy-to-Background Noise Ratio

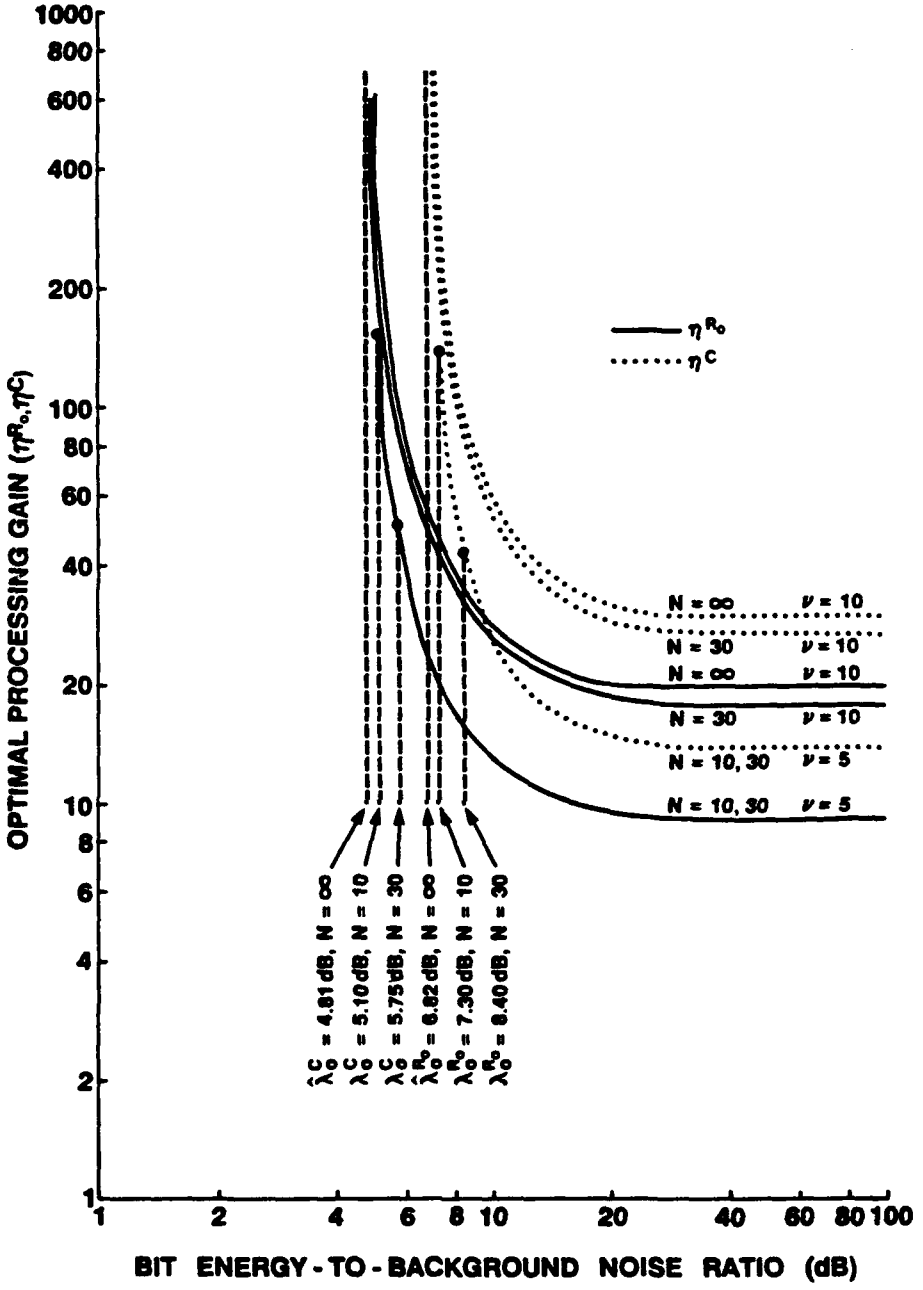


Figure 9: Optimal Processing Gain versus Bit Energy-to-Background Noise Ratio

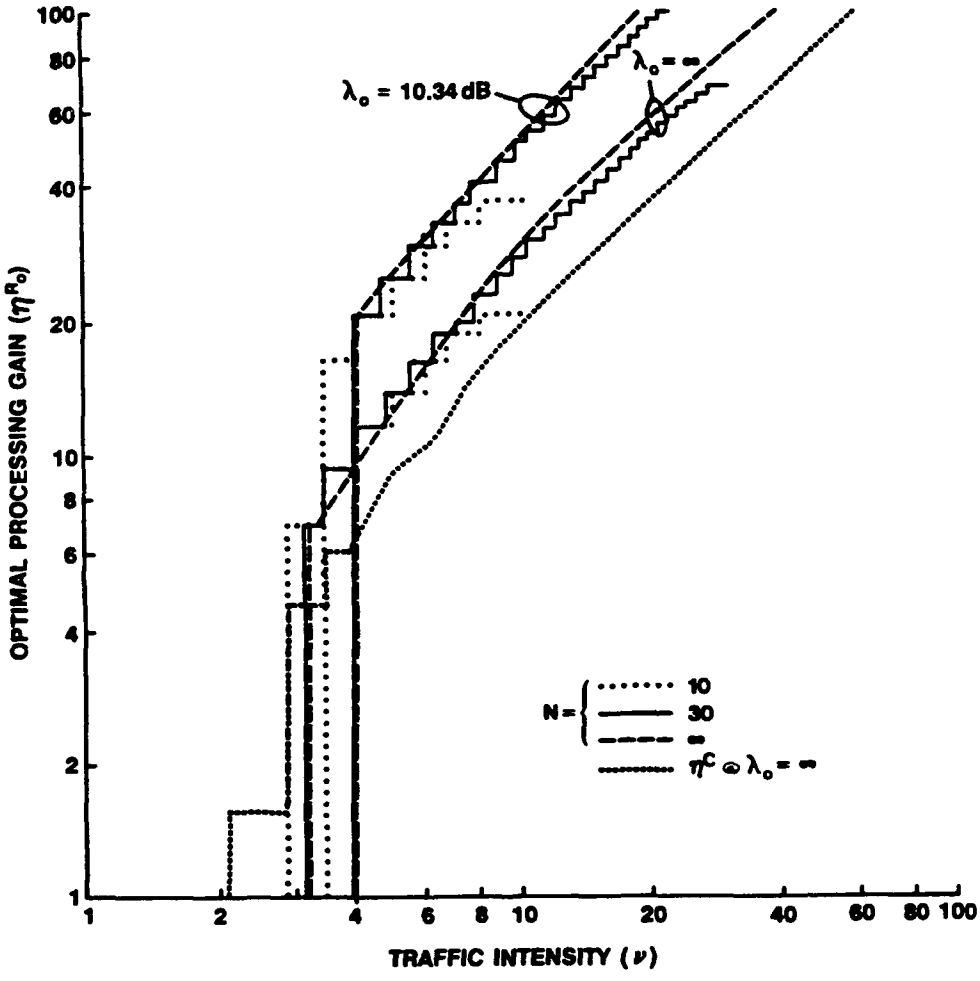


Figure 10: Optimal Processing Gain versus Traffic Intensity for the Cutoff Rate Case (Capacity Case at $\lambda_o = \infty$ is shown for reference)

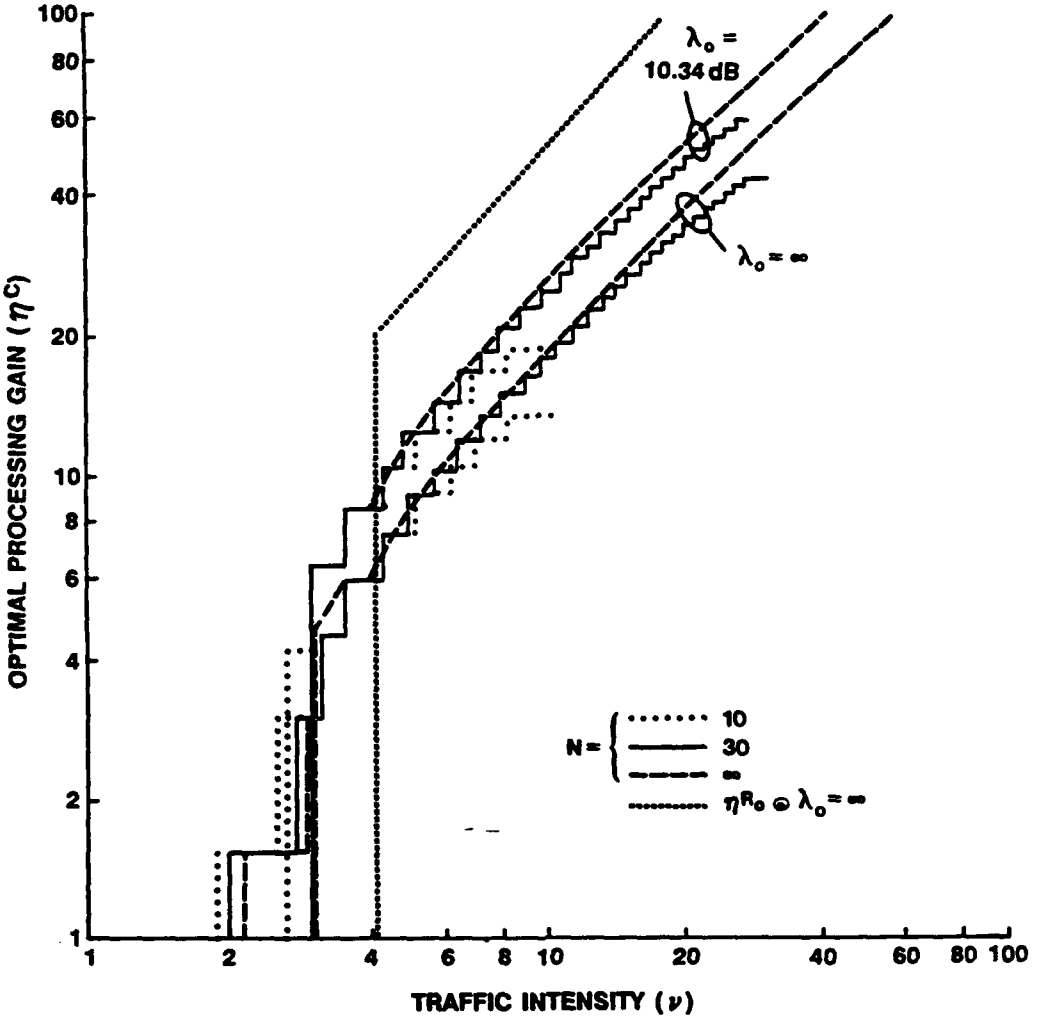


Figure 11: Optimal Processing Gain versus Traffic Intensity for the Capacity Case (Cutoff Rate Case at $\lambda_o = \infty$ is shown for reference)

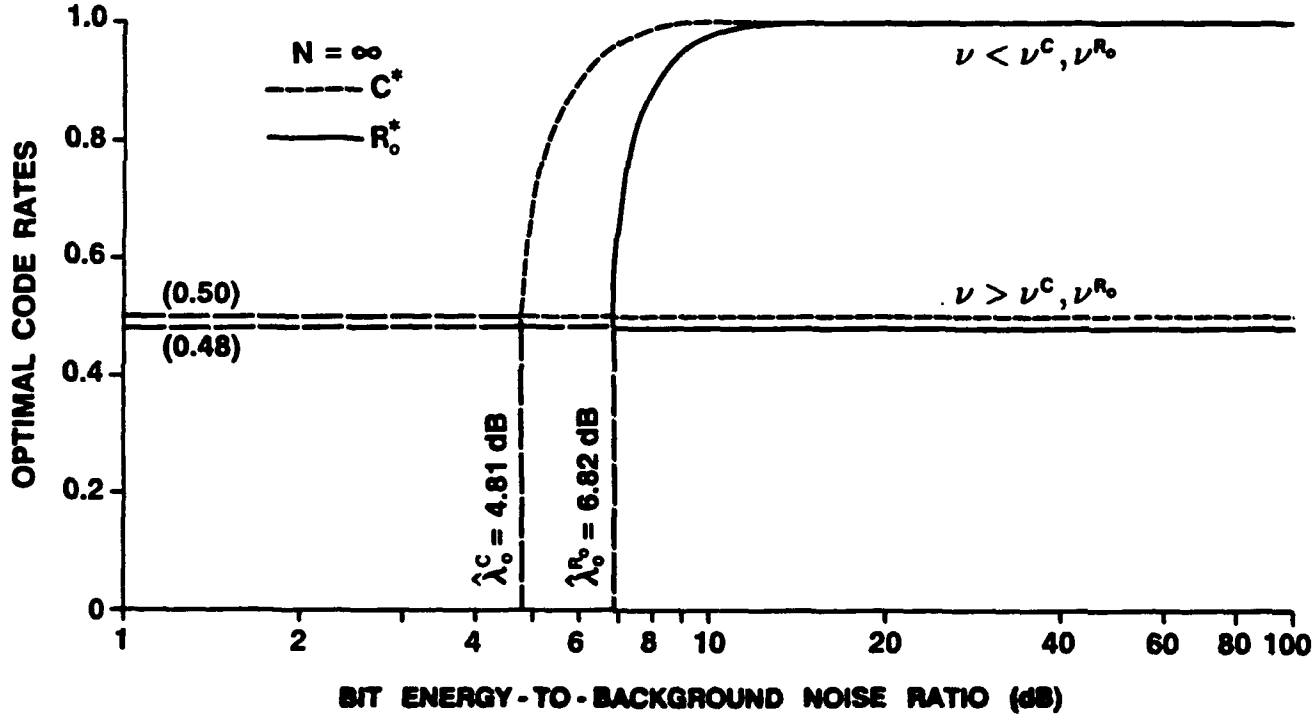


Figure 12: Optimal Code Rate versus Bit Energy-to-Background Noise Ratio

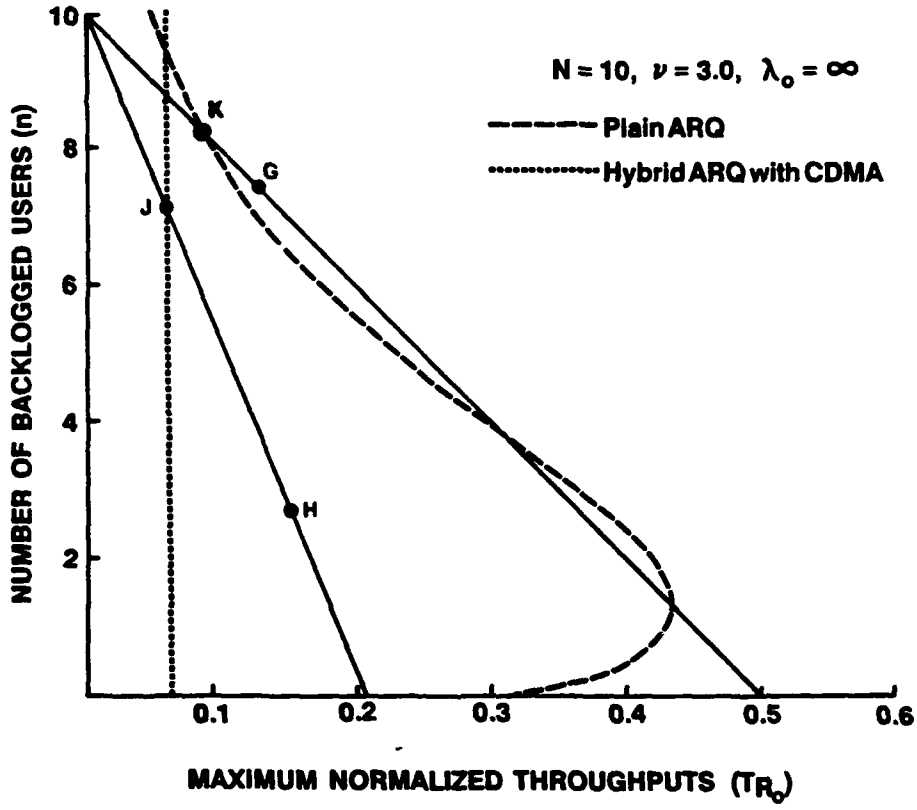


Figure 13: Number of Backlogged Users versus Input/Output Rate Equations for CDMA Network Stability Analysis

DEPARTMENT OF THE ARMY
AIRMICS
115 O'KEEFE BUILDING
GEORGIA INSTITUTE OF TECHNOLOGY
ATLANTA, GA 30332-0800

OFFICAL BUSINESS
PENALTY FOR PRIVATE USE, $300

THIRD CLASS